# The invariant surface glycoprotein ISG75 gene family consists of two main groups in the *Trypanozoon* subgenus

T. TRAN[1,2]*, F. CLAES[1], J. C. DUJARDIN[1] *and* P. BUSCHER[1]

[1] *Institute of Tropical Medicine, Department of Parasitology, Nationalestraat 155, B-2000 Antwerp, Belgium*
[2] *Vrije Universiteit Brussel, Department of Biotechnology, Pleinlaan 2, B-1050 Brussels, Belgium*

(*Received 29 March 2006; revised 6 June 2006; accepted 8 June 2006; first published online 4 September 2006*)

SUMMARY

In *Trypanosoma brucei brucei*, an invariant surface glycoprotein of molecular weight 75 kDa (ISG75) is uniformly distributed over the surface of a trypanosome and is specific for bloodstream-form parasites. For the other taxa of the *Trypanozoon* subgenus no data about this surface molecule are available. Therefore, we investigated the *ISG75* in the genomes of several pathogenic *Trypanozoon* by Southern blot, PCR and RT-PCR and sequence analysis†. This study reveals that (i) all members of the *Trypanozoon* subgenus, i.e. *T. b. brucei*, *T. b. gambiense*, *T. b. rhodesiense*, *T. evansi* and *T. equiperdum*, harbour *ISG75* as multiple gene copies with at least 4–16 copies per genome; (ii) *ISG75* gDNA and cDNA sequences are distributed in 2 groups that share at least 75% and 77% identity respectively; (iii) sequences from both groups are transcribed in all species and subspecies of the *Trypanozoon* subgenus; (iv) the main differences between group I and group II are located in the variable region at the amino-terminus of the putative proteins; (v) however, all the sequences in both groups have some well-conserved features, such as the cysteine residues, an amino-terminal cleavable signal peptide, a single $\alpha$-helix transmembrane domain and a cytoplasmic domain at the carboxy-terminus.

Key words: *Trypanosoma brucei*, *Trypanozoon*, invariant surface glycoprotein (ISG75), molecular characterization.

## INTRODUCTION

The *Trypanozoon* subgenus plays a very important role in trypanosomiasis that affects a variety of mammalian species including man and domestic animals. *Trypanosoma brucei gambiense* and *T. b. rhodesiense* are responsible for sleeping sickness in humans, while *T. b. brucei* and *T. evansi* respectively cause nagana and surra in both domestic animals and game. *T. equiperdum* is believed to cause dourine in equines through sexual contacts.

At its surface, a trypanosome bears a homogenous coat of approximately $10^7$ molecules of variant surface glycoproteins (VSGs) that are highly immunogenic. At any time, a given parasite expresses only 1 type of VSG. This dense layer of VSG dimers acts as a protective umbrella for underlying surface molecules such as the invariant surface glycoproteins (ISGs). Of all the bloodstream-form ISGs that have been characterized so far, including ISG64 (Jackson *et al*. 1993), ISG65 or ISG70 (Ziegelbauer and Overath, 1992), ISG75 (Ziegelbauer *et al*. 1992) and ISG100 (Nolan *et al*. 1997), their functions and structures remain unknown. In a bloodstream-form *T. b. brucei*, there are approximately $5 \times 10^4$ ISG75 molecules evenly distributed over its surface. The deduced ISG75 polypeptide of 523 amino acids is comprised of 4 main regions: an amino-terminal hydrophobic signal sequence (28 amino acids) that is cleaved off yielding a mature protein starting at $E^{29}$; a large hydrophilic extracellular domain; a stretch of hydrophobic residues (20 amino acids) close to the carboxyl-terminus forming a single trans-membrane $\alpha$-helix; and a short hydrophilic domain (29 amino acids) exposed on the cytoplasmic face of the plasma membrane (Ziegelbauer *et al*. 1995). Three *ISG75* sequences of only *T. b. brucei* were revealed by Ziegelbauer and colleagues, and 5 *ISG75* open reading frames have been reported by the *T. brucei* Genome Project (Berriman *et al*. 2005).

This study was designed to extend these observations with the isolation and sequence analysis of *ISG75* gDNA and cDNA in all the pathogenic species and subspecies of *Trypanozoon*, including *T. b. brucei*, *T. b. gambiense*, *T. b. rhodesiense*, *T. evansi* and *T. equiperdum*.

## MATERIALS AND METHODS

### *Trypanosomes and preparation of trypanosomal DNA*

A total of 8 trypanosome strains, of which 5 are cloned populations, were used in this study. Their

* Corresponding author: Institute of Tropical Medicine Antwerp, Department of Parasitology, Nationalestraat 155, B-2000 Antwerp, Belgium. Tel: +32 3 247 65 34. Fax: +327 3 247 63 73. E-mail: thao.tran@vub.ac.be
† Nucleotide sequence data reported in this paper are available in the GeneBank™, EMBL and DDBJ databases under the Accession numbers *DQ200175-DQ200256*.

 © 2006 Cambridge University Press
doi:10.1017/S0031182006000953 Printed in the United Kingdom

https://doi.org/10.1017/S0031182006000953 Published online by Cambridge University Press

Table 1. Description of the *Trypanosoma* (*Trypanozoon*) strains used and the number of ISG75 gDNA and cDNA clones isolated in this study

| Species | Strain | Origin | Host | Year | Number of clones | |
|---|---|---|---|---|---|---|
| | | | | | gDNA | cDNA |
| *T. b. brucei* | AnTat 2.2* | Nigeria | Tsetse | 1970 | 7 | 7 |
| *T. b. gambiense* | LiTat 1.3* | Côte d'Ivoire | Human | 1952 | 6 | 5 |
| *T. b. rhodesiense* | AnTat 12.1 S* | Rwanda | Human | 1972 | 6 | 0 |
| *T. b. rhodesiense* | STIB 850 | Uganda | Human | 1990 | 2 | 6 |
| *T. evansi* | Rotat 1.2* | Indonesia | Buffalo | 1982 | 6 | 6 |
| *T. equiperdum* | Alfort | Unknown | Unknown | Unknown | 7 | 7 |
| *T. equiperdum* | BoTat 1.1.* | Morocco | Horse | 1924 | 7 | 5 |
| *T. equiperdum* | OVI | South Africa | Horse | 1976 | 6 | 0 |

* Cloned.

origins and hosts from which they were isolated are described in Table 1. Trypanosome pellets and total DNA were prepared as described by Claes *et al.* (2003). Quality control of the extracted gDNA was performed by PCR using primer set ESAG6/7 to amplify the transferrin receptor of trypanosomes (Holland *et al.* 2001).

### *Primer design*

A primer set ISG75-1F (TTC TAC GGC CAA GGT AAC GTG), ISG75-2R (GTT CGG GCA CCT GTG ATA GT) was derived from the published 5′-untranslated region and the 3′-untranslated region respectively of the *ISG75* sequence of *T. b. brucei* (Ziegelbauer *et al.* 1995). This primer set was designed to amplify a 1679-bp DNA region including the whole open reading frame of *ISG75* (1572 bp) flanked by the 5′-untranslated region (84 bp) and 3′-untranslated region (22 bp). A second primer set ISG75-3F (AAG GCT GAG AAG GCA AAG GAG), ISG75-2R (GTT CGG GCA CCT GTG ATA G) targeted the 3′-end fragment of *ISG75* (550 bp). The designed primers lacked homology with any other known sequence present in the databases.

### *PCR assays*

To check for the presence of *ISG75*, PCR assays were performed using primer set ISG75-1F, ISG75-2R and ISG75-3F, ISG75-2R with Hotstart Taq polymerase (Qiagen, Germany). The PCR mix contained 1·5 units of Hotstart Taq polymerase, 0·8 $\mu$M of each primer, 0·2 mM dNTPs and 1×Q solution. Conditions were as follows: initial denaturation for 15 min at 95 °C, followed by 35 amplification cycles of 1 min, denaturation at 95 °C, 1 min, primer-template annealing at 55 °C and 1 min elongation at 72 °C. A final extension step was carried out for 5 min at 72 °C. For cloning of full-length *ISG75*, high fidelity Platinum® *Pfx* DNA polymerase (Invitrogen, UK) was employed to isolate the *ISG75* gDNA. The PCR mix contained 1 unit of *Pfx*, 0·4 $\mu$M of each primer, 0·3 mM dTNPs, 1×PCR buffer, 1 mM $Mg^{2+}$. Cycling conditions were as follows: denaturation for 2 min at 94 °C, followed by 35 amplification cycles of 30 sec, denaturation at 94 °C for 30 sec, primer-template annealing at 68 °C and 2 min polymerization at 68 °C. A final extension step was carried out for 10 min at 68 °C.

### *Southern blot and densitometry*

Single restriction endonuclease digestion was performed for each trypanosome strain. Thirty $\mu$l of genomic DNA (approximately 150 ng·$\mu$l$^{-1}$) was digested overnight at 37 °C with 40U of *Hind*III or *Eco*RI (NEB, UK). Twenty $\mu$l of the digested sample and $\lambda$ DNA Molecular Weight Marker II (Eurogentec, Belgium) were resolved by electrophoresis in a 1·5% agarose gel for 17 h at 20 V. Subsequently, DNA from the agarose gel was transferred to a nitrocellulose membrane according to published methods (Sambrook and Russell, 2001).

The 3′-end fragment of *T. b. brucei* AnTat 2.2 obtained via the primer set ISG75-3F, ISG75-2R was used as DNA probe. $\alpha$-$^{32}$P-dCTP labelling of the probe was performed using DecaLabel™ DNA Labeling Kit (MBI Fermentas, Germany). The ProbeQuant™ G-50 Micro Column was used for removal of unincorporated labelled nucleotides (Amersham Biosciences, UK). Hybridization and subsequent washing steps was performed according to the protocols of Sambrook and Russell (2001). The intensity and area of the autogradiographic bands obtained from Southern blotting were quantified by densitometry using Image Master 1D Elite version 3.01 (Pharmacia, UK). Autoradiograms were selected which corresponded closely to an idealized linearity of the signal (Victoir *et al.* 1995).

https://doi.org/10.1017/S0031182006000953 Published online by Cambridge University Press

### *RNA extraction*

Seven *Trypanosoma* strains as described in Table 1 (except *T. b. rhodesiense* AnTat 12.1 S) were used for RNA extraction. Total trypanosome RNA was extracted using the RNAqueous™-Midi Kit (Ambion, UK). Concentration and quality of the isolated RNA were analysed by capillary electrophoresis RNA 6000 Nano Assay, using a RNA 6000 Nano LapChips® (Agilent, UK). To assure that the total RNA in each sample was free of gDNA, a PCR (primer set ISG75-1F, ISG75-2R) was performed using the total RNA as template. One or 2 $\mu$g of the total RNA was mixed with a PCR cocktail containing the components as described for amplification of the full-length *ISG75*.

### *RT-PCR*

The synthesis of first-strand cDNA was carried out using the Omniscript™ Reverse Transcriptase (Qiagen, Germany) according to the manufacturer's manual. The reverse transcription reaction was performed in a Biometra® Trio-block thermocycler at 37 °C for 60 min. For isolation of full-length *ISG75* cDNA, 10 $\mu$l of the RT product were subjected to PCR assay using Platinum® *Pfx* DNA polymerase as described earlier.

### *gDNA and cDNA cloning*

Cloning of the (RT-) PCR products was performed using the Zero Blunt® Topo® PCR Cloning Kit for Sequencing (Invitrogen, UK) according to the manufacturer's manual. To screen for the presence of the desired insert (i.e. 1679-bp *ISG75*) in the cells, PCR was performed using vector T3 and T7 primers flanking the insert. The plasmids containing the desired insert were extracted from the cells using the QIAprep® Miniprep Kits (Qiagen, Germany).

### *DNA sequencing*

The extracted plasmids were sequenced with T3 and T7 primers. The obtained nucleotide sequences were assembled, from which a set of internal primers was designed for the second strand sequencing. The obtained sequences by internal-primer sequencing were assembled and confirmed with the sequences that were acquired by the T3–T7 sequencing. Sequence analysis was performed using DNAman (Lynnon Soft, Australia) and publicly available softwares, including InterProScan for topology prediction (http://www.ebi.ac.uk/InterProScan/), SignalP for signal peptide prediction (http://www.cbs.dtu.dk/services/SignalP/), and NetNGlyc for N-linked glycosylation prediction (http://www.cbs.dtu.dk/services/NetNGlyc/).

## RESULTS

### *Sequence analysis of full-length ISG75 gDNA and cDNA clones*

As there were 3 *ISG75* sequences of only *T. b. brucei* reported, we intended to investigate the presence and transcription of *ISG75* in all species and subspecies of the *Trypanozoon*. A 1679-bp PCR product of the full-length *ISG75* was isolated in all the tested strains of the *Trypanozoon* subgenus with either Hotstart *Taq* or *Pfx* DNA polymerase. The same result was obtained with RT-PCR in all these strains. Therefore, it shows that *ISG75* is present and transcribed in all species and subspecies of the *Trypanozoon*.

In total, 10–14 gDNA and cDNA clones isolated from each strain were sequenced (Table 1). The isolated gDNA and cDNA sequences show that multiple copies of *ISG75* are present and transcribed in the genomes of all the strains. A nucleotide similarity tree constructed from 83 sequences divides the *ISG75* cDNA and gDNA of the examined strains into 2 groups (Group I and Group II) with 77% and 75% identity respectively. For clarity purposes, here we present the sequence alignment of all the gDNA as 1 similarity tree (Fig. 1), and of all the cDNA as another (Fig. 2). For either gDNA or cDNA, highly conserved sequences ($\geqslant$90% or $\geqslant$94% identity) in Group I are present across different (sub)species of *T. b. brucei* AnTat 12.1, *T. b. gambiense* LiTat 1.3, *T. b. rhodesiense* AnTat 12.1 S and STIB 850 HR, *T. evansi* RoTat 1.2 and *T. equiperdum* Alfort. Likewise, Group II harbours highly similar gDNA and cDNA sequences among all strains ($\geqslant$93% or $\geqslant$92% identity). Out of 5 DNA sequences of *T. b. brucei* reported by the genome project, 2 (Tb927.5.390, Tb927.5.370) are clustered in Group I and 2 (Tb927.5.350, Tb927.5.400) belong to Group II in our analysis. However, Tb927.5.380 does not fall into either groups, and has 52% identity with Group I and Group II (Fig. 1). Of the 3 *ISG75* sequences identified by Ziegelbauer and colleagues, M86710 and L07866 belong to Group I; and L07867 belongs to Group II in this gene family (Fig. 2). Furthermore, *in silico* and experimental resstriction analysis of the iolated sequences with *Mlu*I revealed that all the gDNA and cDNA sequences in Group I contain a unique *Mlu*I site, that is absent in Group II. The RT-PCR of *T. equiperdum* OVI was negative; hence no ISG75 cDNA of *T. equiperdum* OVI were obtained. cDNA sequences of only Group II were obtained in *T. equiperdum* BoTat 1.1. However, PCR of *T. equiperdum* Botat 1.1 and OVI gDNA using Group I- and Group II-specific primers and *Mlu*I restriction digestion were positive for both groups indicating that ISG75 gDNA sequences of Group I and Group II are also present in these two strains.

https://doi.org/10.1017/S0031182006000953 Published online by Cambridge University Press

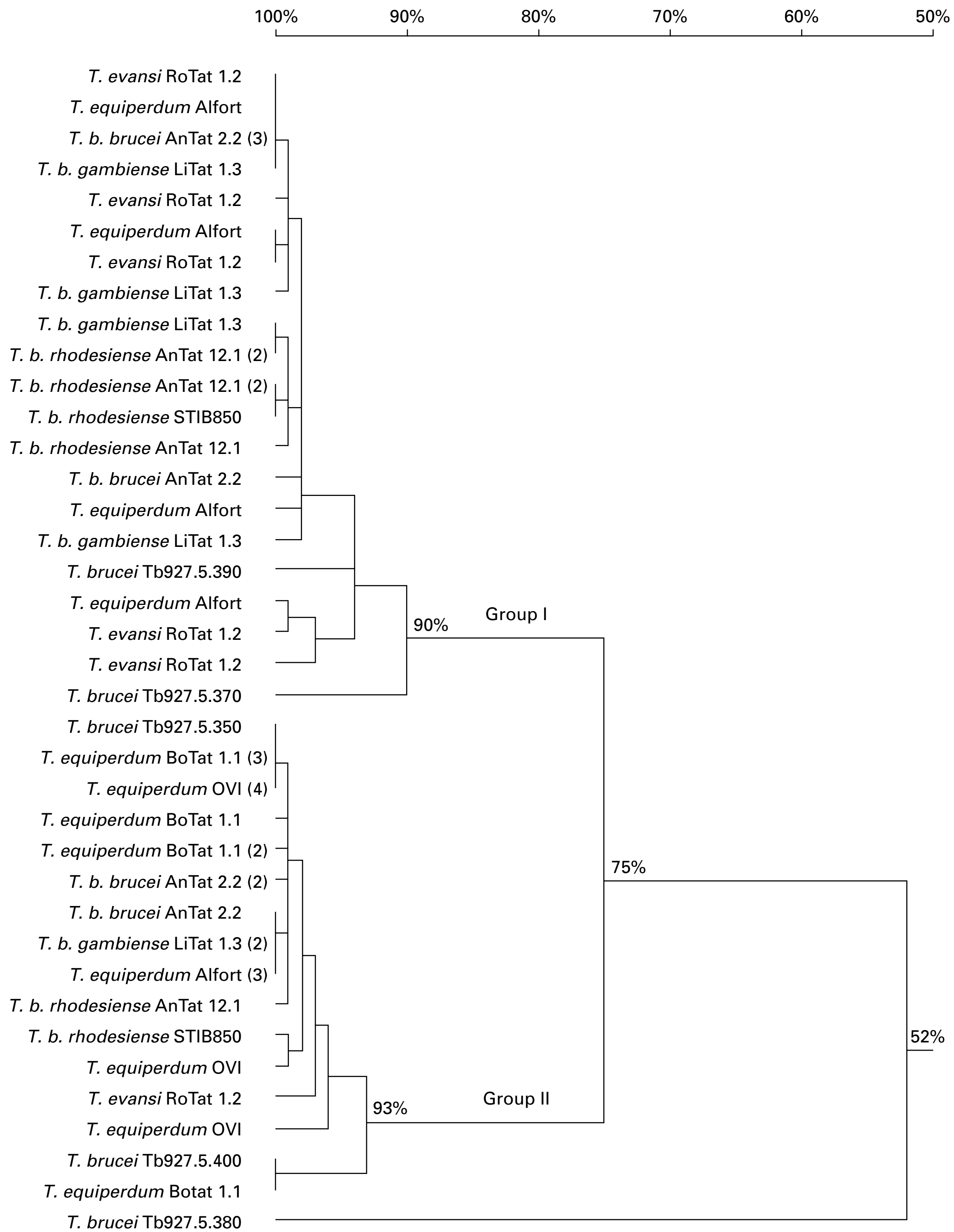


Fig. 1. *ISG75* gDNA similarity tree, constructed with DNAman. Numbers in parentheses indicate the number of identical sequences that were isolated from the *E. coli* clones of the particular PCR product.

All the cDNA and gDNA were translated into amino acid sequences and multiple alignment of the putative amino acid sequences distinguished Group I and Group II based on the following. Firstly, all the sequences in Group II lack 1 amino acid residue at position 157 as compared to Group I (S or L). As a result, Group I sequences have 523 residues (except 2 gDNA sequences of *T. b. gambiense* that had 520 residues, lacking residues $A^{147}E^{148}K^{149}$); Group II sequences have 522. Secondly, all Group II sequences have a potential N-linked glycosylation site at position 134 (tripeptide NSS) and most of the Group II sequences have a second site at position 115 (tripeptide NRT); whereas all Group I sequences have only 1 potential N-linked glycosylation site at position 134 (tripeptide NAS) (Fig. 3).

https://doi.org/10.1017/S0031182006000953 Published online by Cambridge University Press

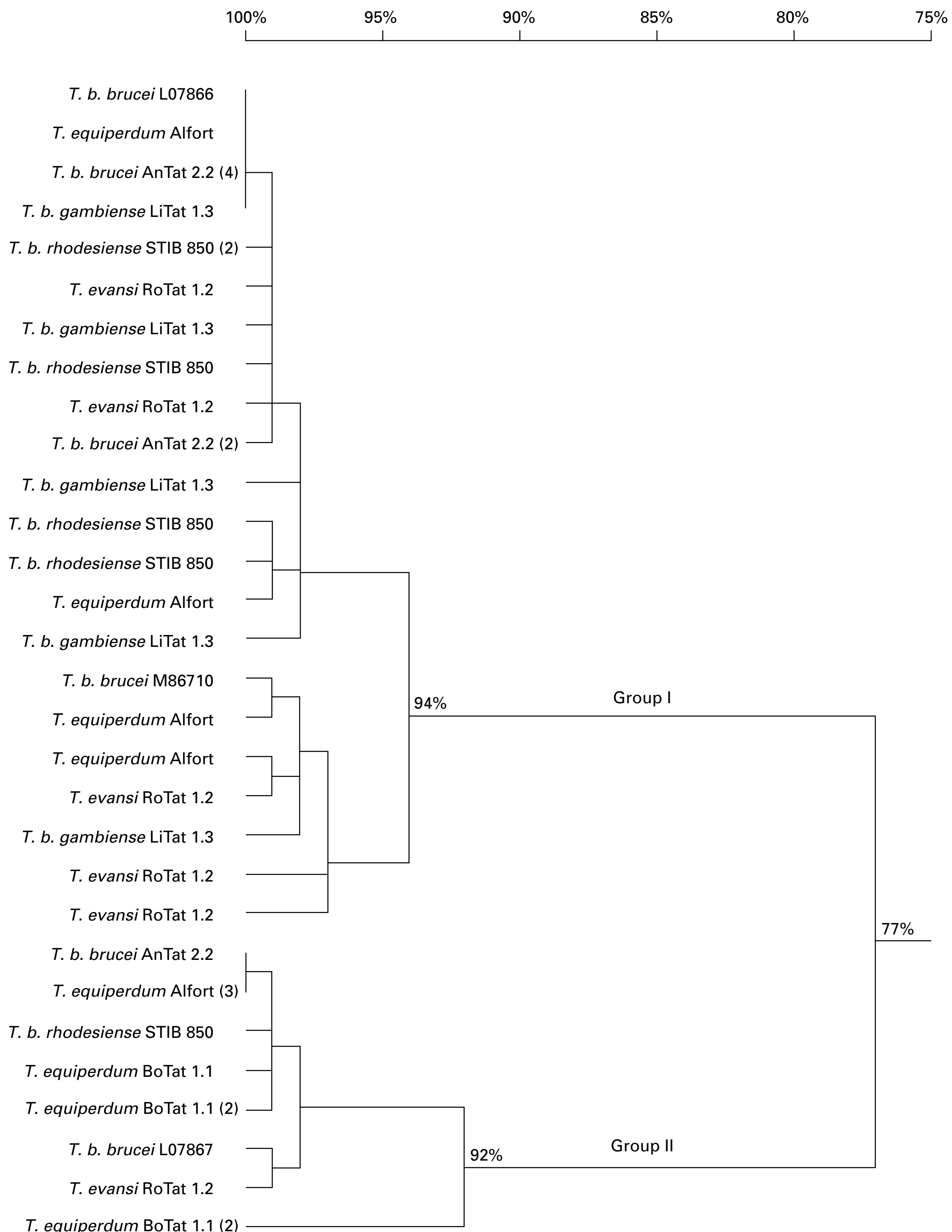


Fig. 2. *ISG75* cDNA similarity tree, constructed with DNAman. Numbers in parentheses indicate the number of identical sequences that were isolated from the *E. coli* clones of the particular RT-PCR product.

Multiple alignment and hydrophobicity analysis of the putative amino acid sequences demonstrate that ISG75 contains 2 conserved hydrophobic regions: one 28-amino acid segment at the amino terminus indicates a potential signal peptide; and one 21-amino acid region close to the carboxy end, suggests a single $\alpha$-helix trans-membrane domain of the protein (Fig. 3). The 34-residue segment at the carboxy-terminus is predicted to be a cytoplasmic domain. The signal peptide cleavage site is predicted

https://doi.org/10.1017/S0031182006000953 Published online by Cambridge University Press

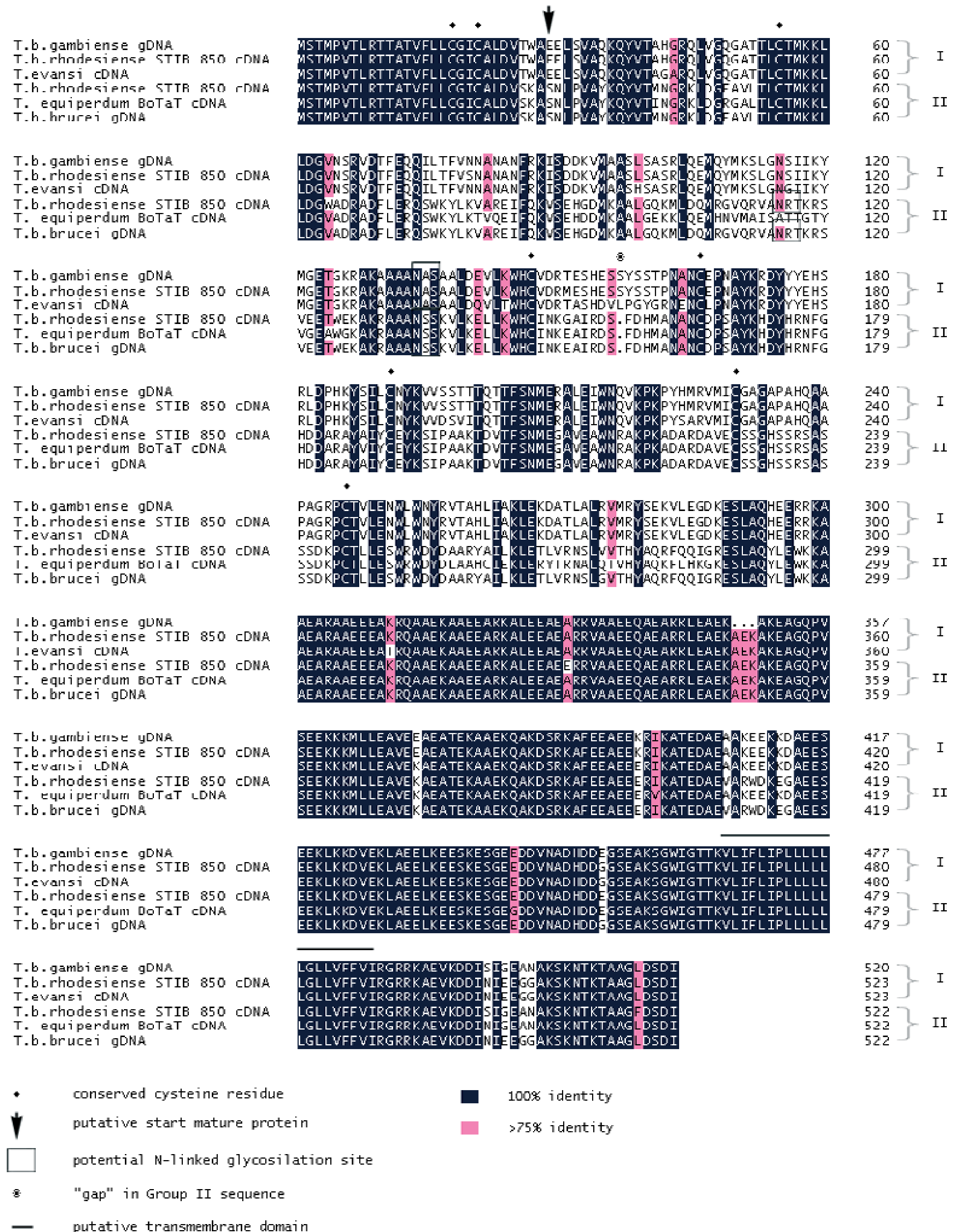


Fig. 3. Multiple alignment of some of the putative amino acid sequences, constructed with DNAman, highlighting the variable and conserved regions between Group I and group II ISG75.

to be between residue 28 and 29, hence the mature protein potentially starts at residue $E^{29}$ for most of the Group I sequences or $S^{29}$ for most of the Group II sequences. A single $\alpha$-helix trans-membrane domain is predicted at aa 469–491 for Group I sequences, and at aa 468–490 for Group II sequences with 100% sequence identity among all sequences of both groups. A large region between residue 29 and

https://doi.org/10.1017/S0031182006000953 Published online by Cambridge University Press

467 might serve as an extracellular domain of the protein. The region immediately downstream of the signal sequence, from amino acid 29 to approximately amino acid 300, is heterogeneous; whereas roughly 200 amino acids thereafter are highly conserved. The overall predicted topology of the sequences in both groups is consistent with that of *T. b. brucei* isolated by Ziegelbauer *et al.* (1992).

In addition to the conserved domains of ISG75 in both Groups I and II, the variable region also shares particularly well-conserved features in all the sequences. These include an N-linked glycosylation site at position 134 as described earlier and 6 conserved cysteines that are located at identical positions in the variable region (Fig. 3). Together with 2 cysteines located in the signal peptide, there are overall 8 conserved cysteines in all the ISG75 (except overall 9 cysteines in 3 of the *T. equiperdum* BoTat 1.1 sequences).

### *Estimation of ISG75 copy number*

Complementary to the isolation of ISG75 gDNA and cDNA sequences in all species and subspecies of the *Trypanozoon*, Southern blot and densitometry were performed for estimation of the copy number of the ISG75 in the genomes of these strains. It was verified by restriction digest and sequence analysis that neither *Hind*III nor *Eco*RI restriction sites were present in the probe region. Southern blot of gDNA digested with the restriction endonuclease *Hind*III or *Eco*RI showed multiple bands in all tested strains. Densitometrical scanning was performed with the bands obtained from the Southern blot for estimation of its copy number. In each lane, the band with lowest intensity was considered as at least 1 gene copy. The results indicate that *ISG75* copy number fluctuates among strains of the *Trypanozoon*, ranging from at least 4 to 16 copies in each strain.

## DISCUSSION

This study confirms that all *Trypanozoon* species and subspecies harbour and transcribe multiple *ISG75* sequences. Interestingly, regardless of the strains, all the gDNA and cDNA cluster into 2 groups (Group I and Group II). These sequences have high similarities with the *ISG75* sequences of *T. b. brucei* reported in the databank. All 3 ISG75 sequences identified by Ziegelbauer and colleagues, and 4 out of 5 sequences reported in the *T. brucei* genome project fall into Group I and Group II in our analysis. Previous observations in terms of glycosylation site and variations/substitutions of residues in the cytoplasmic and signal peptide domains correlate well with our findings in the two groups (Ziegelbauer *et al.* 1995). Therefore, this study in a sense is an extension to all members of the *Trypanozoon* subgenus of the previously reported results on *T. b. brucei ISG75*. Moreover, this study is the first to identify the 2 main groups of the *ISG75* gene family in the *Trypanozoon* subgenus. In agreement with the previous work on *T. b. brucei* and the high correlation between the Southern blot and gene cloning results in this study, *ISG75* is present as multiple copies in the genomes of all the taxa of the *Trypanozoon* subgenus.

The third-group sequence Tb927.5.380 identified in the *T. brucei's* genome project was not isolated in our experiment. As our primers were designed based on the untranslated regions of ISG75, this third group sequence either has different untranslated regions or is absent in the tested strains. Furthermore, it has little similarity (approximately 33% identity) with either Group I or Group II at amino acid level. The only similarities between this sequence and the rest of the ISG75 members are a predicted signal peptide, a predicted transmembrane domain and positions of 6 (out of 7) cysteines in the mature putative polypeptide. In addition, it has 3 deletions (total of 15 amino acids) and 4 insertions (total of 44 amino acids) when aligned to the other ISG75 sequences. Therefore, it is approximately 29–30 residues longer than all other ISG75 open reading frames available in the databank. Further research is needed to reveal the general presence and expression of this sequence in the *Trypanozoon* group and its relationship with Group I and Group II ISG75.

*In silico* analysis based on preliminary genome sequences of *T. congolense* and *T. vivax* (http://www.sanger.ac.uk/Projects) shows that there is no sequence identity between the full-length ISG75 and *T. congolense* and *T. vivax*'s sequences. It suggests that *ISG75* is probably specific for *Trypanozoon*. Therefore, this may have implications on possible differences in surface architecture or host survival mechanisms between the *Trypanozoon* and the other subgenera of *Trypanosoma*.

It is a common feature in Kinetoplastids that genes are transcribed as polycistronic mRNA. The mature mRNA is generated by addition of a miniexon at the 5′-end of the open reading frame and a polyadenylation tail at the 3′-end (Borst, 1986; Agabian, 1990). In Trypanosomatids, several of their characterized genes belong to tandemly arrayed multigene families, including calmodulin (Tschudi *et al.* 1985), tubulin (Thomashow *et al.* 1983; Seebeck *et al.* 1983) and trypanosome hexose transporters (Bringaud and Baltz, 1994). The *ISG75* gene family is of no exception and shows an interesting genetic diversity. Highly similar sequence members (at least 92% identity in Group I and 90% identity in Group II at amino acid level) are present across different strains, while the variations between Group I and Group II (67% identity at amino acid level) are mainly situated in discrete segments at the variable regions. In addition, the difference in length of most of the

https://doi.org/10.1017/S0031182006000953 Published online by Cambridge University Press

putative amino acid sequences between Groups I and II is by only 1 residue at the same position.

During the course of their evolution, the *Trypanozoon* have kept the transmembrane and cytoplasmic domain of the ISG75 highly conserved; perhaps because they are well protected in the parasite's membrane and not prone to the host immune system, or perhaps because they have some particular function. However, the parasite allows genetic variability in the extracellular domain that is relatively more exposed to the host environment, hence is able to adapt itself to survival in the mammalian hosts. It is noted that as ISG75 is smaller than the VSG, it is probably hidden in the VSG coat on the surface (Overath *et al*. 1994). Therefore, it remains unclear why there is a variable extracellular region among all the sequences and why 2 groups of *ISG75* sequences are transcribed. It may be hypothesized that different protein sequences from both groups are required for adaptation to different mammalian host species as has been proposed for the transferrin receptor ESAG6/7 that has different affinity for types of transferrins (Isobe *et al*. 2003). Also it has been reported that the trypanosome hexose transporter THT1 and THT2 that have 6 and 5 direct repeats in *tht* loci respectively and are developmentally regulated for glucose availability in different hosts (Bringaud and Baltz, 1993, 1994). One could imagine that ISG75, being smaller and less abundant than the VSGs, may play a role as a receptor of some particular ligand for signalling between the host's environment and the parasite or among the parasites themselves. Furthermore, it is possible that 2 molecules of Groups I and II are necessary for the formation of a functional dimeric structure. Since several surface molecules of the trypanosome are known to be dimers, ISG75 might also be either a homodimer formed by either Group I or Group II sequences, like VSG (Freymann *et al*. 1984); or a heterodimer formed by 2 monomers from Group I and Group II, like the combination ESAG6 and ESAG7 for the transferrin receptor (Salmon *et al*. 1994).

In conclusion, ISG75 is 'invariant' in terms of its presence and transcription in all the *Trypanozoon* species and subspecies. However, there is a certain degree of genetic variability among different sequences, especially between Group I and Group II, as it is typical of the *Trypanosomatids*, whose genomes appear to have multiple gene copies that play important roles in their successful parasitic life.

We are grateful to Dr Xu Ying and Dr Jan Van Den Abbeele (Institute of Tropical Medicine, Antwerp) for their practical and scientific assistance. We greatly appreciate the critical comments from Dr Henri De Greve (Free University of Brussels) on the manuscript. This study received financial support from the International Livestock Research Institute (ILRI) in Nairobi, Kenya and a private funding.

REFERENCES

**Agabian, N.** (1990). Transplicing of nuclear pre-mRNAs. *Cell* **61**, 1157–1160.

**Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Bohme, U., Hannick, L., Aslett, M. A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U. C., Arrowsmith, C., Atkin, R. J., Barron, A. J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T. J., Churcher, C., Clark, L. N., Corton, C. H., Cronin, A., Davies, R. M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M. C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B. R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A. X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., MacLeod, A., Mooney, P. J., Moule, S., Martin, D. M. A., Morgan, G. W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C. S., Peterson, J., Quail, M. A., Rabbinowitsch, E., Rajandream, M. A., Reitter, C., Salzberg, S. L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A. J., Tallon, L., Turner, C. M., Tait, A., Tivey, A. R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M. D., Embley, T. M., Gull, K., Ullu, E., Barry, J. D., Fairlamb, A. H., Opperdoes, F., Barrell, B. G., Donelson, J. E., Hall, N., Fraser, C. M., Melville, S. E. and El Sayed, N. M.** (2005). The genome of the African Trypanosome *Trypanosoma brucei*. *Science* **309**, 416–422.

**Borst, P.** (1986). Discontinuous transcription and antigenic variation in trypanosomes. *Annual Review of Biochemistry* **55**, 701–732.

**Bringaud, F. and Baltz, T.** (1993). Differential regulation of two distinct families of glucose transporter genes in *Trypanosoma brucei*. *Molecular and Cellular Biology* **13**, 1146–1154.

**Bringaud, F. and Baltz, T.** (1994). African trypanosome glucose transporter genes: organization and evolution of a multigene family. *Molecular Biology and Evolution* **11**, 220–230.

**Claes, F., Agbo, E. E. C., Radwanska, M., Baltz, T., De Waal, D. T., Goddeeris, B. M. and Büscher, P.** (2003). How does *Trypanosoma equiperdum* fit into the *Trypanozoon* group? A cluster analysis by Random Amplified Polymorphic DNA (RAPD) and Multiplex Endonuclease Genotyping Approach (MEGA). *Parasitology* **126**, 425–431.

**Freymann, D. M., Metcalf, P., Turner, M. and Wiley, D. C.** (1984). 6 A-resolution X-ray structure of a variable surface glycoprotein from *Trypanosoma brucei*. *Nature, London* **311**, 167–169.

**Holland, W., Claes, F., My, L. N., Thanh, T., Tam, P. T., Verloo, D., Büscher, P., Goddeeris, B. M. and Vercruysse, J.** (2001). A comparative evaluation of parasitological tests and a PCR for *Trypanosoma evansi* diagnosis in experimentally infected water buffaloes. *Veterinary Parasitology* **97**, 23–33.

**Isobe, T., Holmes, E. C. and Rudenko, G.** (2003). The transferrin receptor genes of *Trypanosoma equiperdum*

are less diverse in their transferrin binding site than those of the broad-host range. *Journal of Molecular Evolution* **56**, 377–386.

**Jackson, D. G., Windle, H. J. and Voorheis, H. P.** (1993). The identification, purification, and characterization of two invariant surface glycoproteins located beneath the surface coat barrier of bloodstream forms of *Trypanosoma brucei*. *Journal of Biological Chemistry* **268**, 8085–8095.

**Lanham, S. M. and Godfrey, D. G.** (1970). Isolation of salivarian trypanosomes from man and other mammals using DEAE-cellulose. *Experimental Parasitology* **28**, 521–534.

**Nolan, D. P., Jackson, D. G., Windle, H. J., Pays, A., Geuskens, M., Michel, A., Voorheis, H. P. and Pays, E.** (1997). Characterization of a novel, stage-specific, invariant surface protein in *Trypanosoma brucei* containing an internal, serinerich, repetitive motif. *Journal of Biological Chemistry* **272**, 29212–29221.

**Overath, P., Chaudhri, M., Steverding, D. and Ziegelbauer, K.** (1994). Invariant surface proteins in bloodstream forms of *Trypanosoma brucei*. *Parasitology Today* **10**, 53–58.

**Salmon, D., Geuskens, M., Hanocq, F., Hanocq-Quertier, J., Nolan, D., Ruben, L. and Pays, E.** (1994). A novel heterodimeric transferrin receptor encoded by a pair of VSG expression site-associated genes in *T. brucei*. *Cell* **78**, 75–86.

**Sambrook, J. and Russell, D.** (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press. New York.

**Seebeck, T., Wittaker, P. A., Imboden, M. A., Hardman, N. and Braun, R.** (1983). Tubulin genes of *Trypanosoma brucei*: a tightly cluster family of alternating genes. *Proceedings of the National Academy of Sciences, USA* **80**, 4634–4638.

**Thomashow, L. S., Milhausen, M., Rutter, W. J. and Agabian, N.** (1983). Tubulin genes are tandemly linked and clustered in the genomes of *Trypanosoma brucei*. *Cell* **32**, 35–43.

**Tschudi, C., Young, A. S., Ruben, L., Patton, C. L. and Richards, F. F.** (1985). Calmodulin genes in trypanosomes are tandemly repeated and produce multiple mRNA with common 5′-end leader sequence. *Proceedings of the National Academy of Sciences, USA* **82**, 3998–4002.

**Victoir, K., Dujardin, J. C., de Doncker, S., Barker, D. C., Arevalo, J., Hamers, R. and Le Ray, D.** (1995). Plasticity of gp63 gene organization in *Leishmania* (*Viannia*) *braziliensis* and *Leishmania* (*Viannia*) *peruviana*. *Parasitology* **111**, 265–273.

**Ziegelbauer, K., Multhaup, G. and Overath, P.** (1992). Molecular characterization of two invariant surface glycoproteins specific for the bloodstream stage of *Trypanosoma brucei*. *Journal of Biological Chemistry* **267**, 10797–10803.

**Ziegelbauer, K. and Overath, P.** (1992). Identification of invariant surface glycoproteins in the bloodstream stage of *Trypanosoma brucei*. *Journal of Biological Chemistry* **267**, 10791–10796.

**Ziegelbauer, K., Rudenko, G., Kieft, R. and Overath, P.** (1995). Genomic organization of an invariant surface glycoprotein gene family of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* **69**, 53–63.

https://doi.org/10.1017/S0031182006000953 Published online by Cambridge University Press